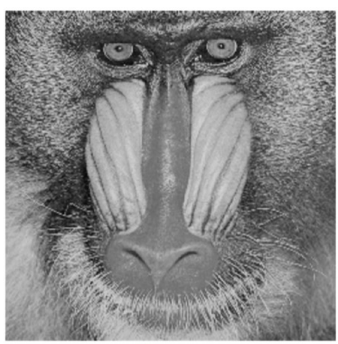
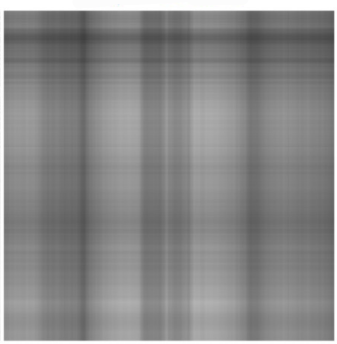
Applications

1.) Image Processing / Compression

Suppose an image is made of $m \times n$ pixels. Each pixel can be one of x colors. We can describe these pixels with an $m \times n$ matrix A. This is contains $m \cdot n$ data points. The SVD k-rank approximation A_k uses much less data and is used to compress the data

Ex:
A gray scale image.
Each entry has a value btwn 0 & 255
(measures brightness)



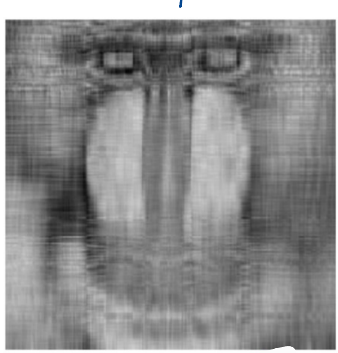


A is 512×512
$0 \leq a_{ij} \leq 1 \quad \forall i,j$

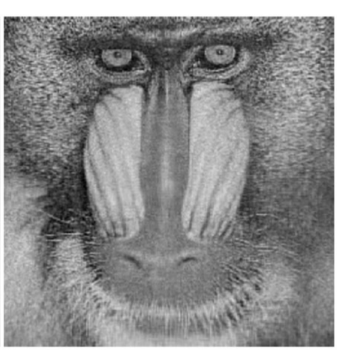
$\Rightarrow 512^2 = 262{,}144$ data points

$A_1 = \sigma_1 u_1 v_1^T$
($m \times 1$, $1 \times n$)

$\Rightarrow m+n+1 = 1025$ data points



A_{10}



A_{50}

$\Rightarrow 50(m+n+1) = 51{,}250$ data points

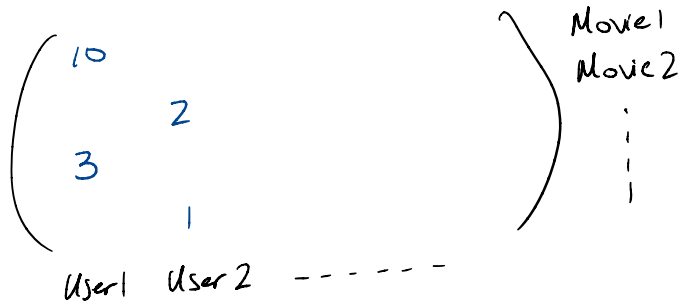
2. Recommendation Systems

Netflix Prize (2007)

Develop algorithm to predict user ratings of movies.

Let A be a (huge) matrix where

$$a_{ij} = \text{rating of movie } j \text{ by user } i$$

$A =$ 

This matrix has mostly empty entries.

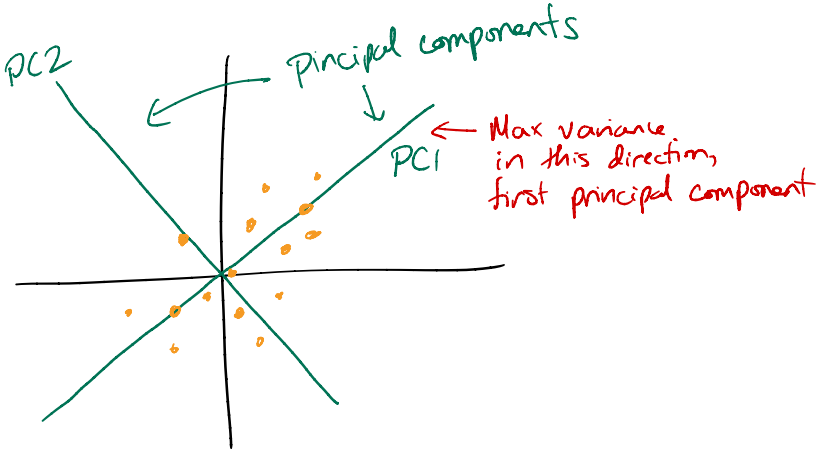
So we first must populate the entries using some heuristic

Next, assuming A has a good low rank approximation (most singular values are close to 0), We can use low rank approx to estimate user ratings.

This is what was done to win the Netflix prize. See links on canvas for more details.

3. Principal Component Analysis (PCA)

PCA is used in statistics and data analysis.
It uses best fit subspaces to understand the variance in a set of data. In particular, it identifies orthogonal axes (principal components) that capture the maximum variance of data.



Suppose $x_1, \dots, x_m \in \mathbb{R}^n$ are m samples containing n measurements.

Ex: Record exam 1 and exam 2 scores for all students in 21-241

$$x_1 = \begin{bmatrix} 92 \\ 83 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 72 \\ 100 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 100 \\ 87 \end{bmatrix}, \quad \dots$$

Center the data about the origin (make each average measurement 0) by subtracting the i^{th} entry of each vector by the average i^{th} measurement

Ex: Avg Score of Exam 1 is 88
Avg Score of Exam 2 is 90

$\Rightarrow x_1 = \begin{bmatrix} 4 \\ -7 \end{bmatrix}, x_2 = \begin{bmatrix} -16 \\ 10 \end{bmatrix}, x_3 = \begin{bmatrix} 12 \\ -3 \end{bmatrix}$ Avg Scores now 0

Let $A = \begin{bmatrix} x_1^T \\ \vdots \\ x_m^T \end{bmatrix}$ contain the centered data.

The sample variance of $x_1, \dots, x_m \in \mathbb{R}^n$ is given by the matrix $\frac{A^TA}{n-1}$ (Covariance matrix)

Note: The (i,i)-entry of $\frac{A^TA}{n-1}$ is the sample variance of the i^{th} entries of the data vectors

It turns out that the eigenvector of A^TA corresponding to the i^{th} largest eigenvalue provides the i^{th} principal component.

So if $A = U\Sigma V^T$ is an SVD, then the i^{th} column of V gives the i^{th} principal component
Moreover, we can use best fit subspaces given by the singular vectors / principal components to approximate the data and provide scatter plots (See recitation)